

ABSTRACT

This case study centers on the detailed analysis of an unknown DNA fragment. The unknown DNA sequence was found to originate from the genome of *Helicobacter Pylori*, using BLASTp. By using bioinformatic platforms such as NCBI ORFfinder, BLASTp, UniProt, and phylogenetic tree construction tools, we research the genetic properties of *H. pylori* and highlight the biological implications of a major **virulence factor** called **arginase**.

INTRODUCTION

Name:	Helicobacter pylori
Organism:	Bacterium
Category:	Gram negative
Shape:	Helical and S-shaped
Mobility:	Flagella & chemotactic action
Incubation time:	3 days
Transmission:	Oral - Oral, Fecal - Oral (through water or food consumption)

Table 1: General information of the organism *Helicobacter Pylori*.

In this case study, the unknown DNA sequence was found to originate from the genome of *Helicobacter Pylori*. *Helicobacter pylori* is a **spiral-shaped, gram-negative bacterium** that infects the lining of the **human stomach** and is one of the most common **chronic bacterial infections** worldwide, affecting more than half the global population. It is typically acquired during **childhood**, especially in areas with poor sanitation or crowded living conditions, and is often transmitted through person-to-person contact or contaminated food and water. While many people with *H. pylori* infection have no symptoms, the bacterium can cause chronic gastritis, peptic ulcers, and, in rare cases, leads to stomach cancer or MALT lymphoma.

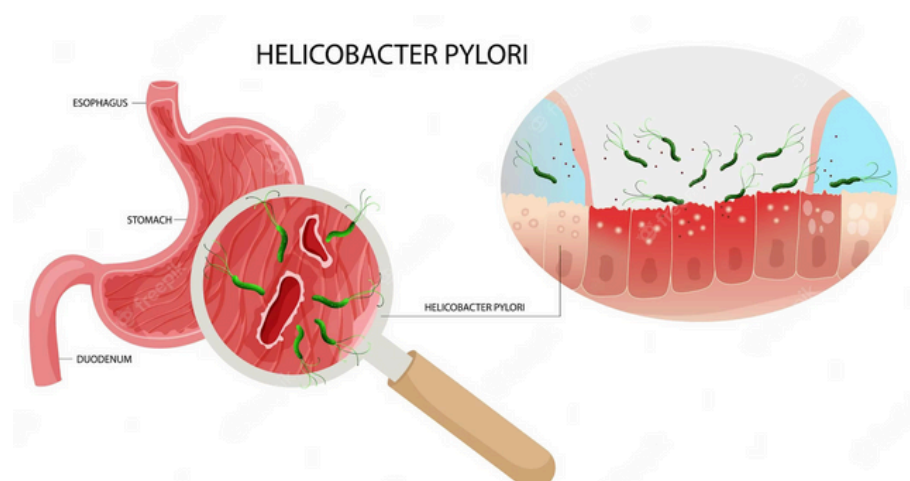


Figure 1: Location of *Helicobacter Pylori* in the gut

MATERIALS AND METHODS

Analysis of DNA fragments

NCBI ORF finder website is used to find the ORFs

- Settings used: standard settings of the site.

BLASTP is done to compare the protein sequence against a protein database to find similar sequences.

- Settings used: standard settings, the database selected: UniProtKB/Swiss-Prot(swissprot).

Literature research

To have a deeper understanding of *Helicobacter Pylori*, a literature study was done in respect to the virulence of the bacteria, the symptoms of infected individuals and the potential therapeutic effects Arginase might have.

Comparative protein study

The NCBI website was used to search for similar proteins. 20 similar protein sequences are selected. Selection requirements:

- Do not 100% Per. Ident.
- Do not have same AA length

Clustral Omega webserver is used to form a Guide tree and phylogenetic tree, to later use in iTOL. The tool called "iTOL (interactive tree of life)" is used to build a phylogenetic tree. The colour coding is done based on taxonomy (the three domains), this information was found on NCBI.

Detailed protein analysis

To start, the sequence of arginase was put in InterPro to identify domains, biological processes and functions. Furthermore, the STRING database was to visualize the protein network of arginase. At last to further visualize the protein in 2D and 3D PDB and AlphaFold were used.

RESULTS

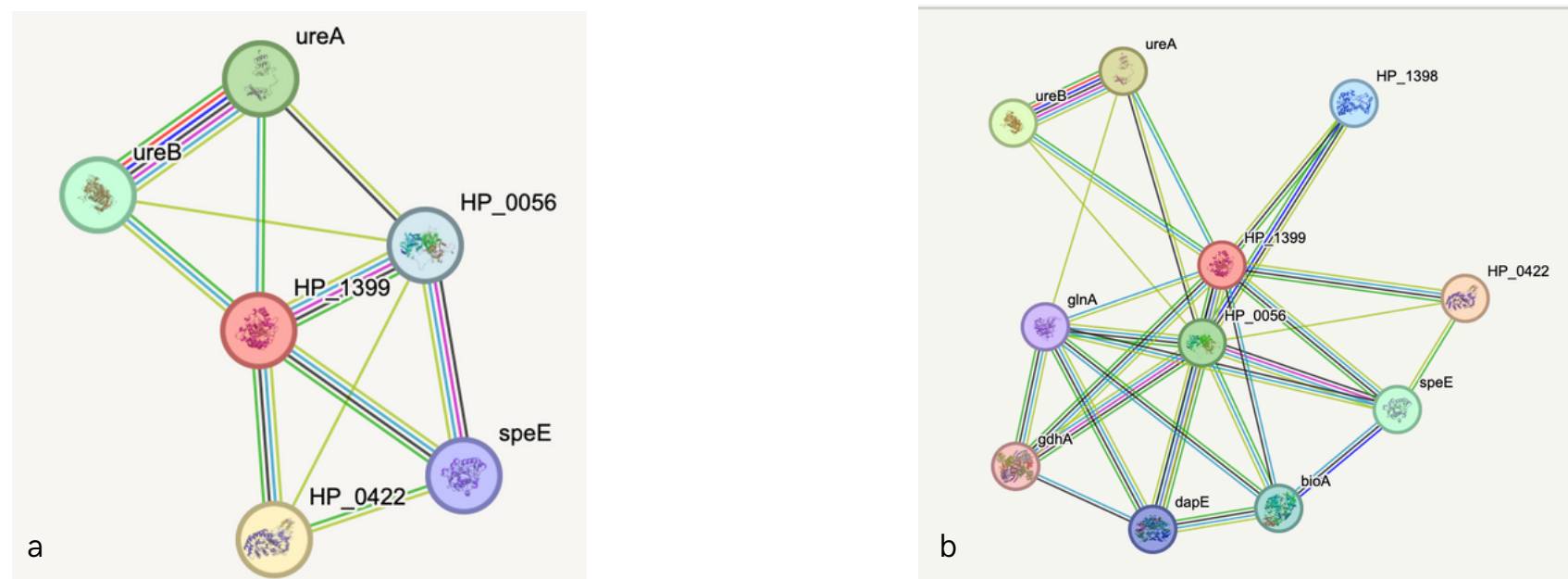


Figure 2: Result of STRING network of the protein Arginase. Interaction network of Arginase (b) and the smaller interactions (a). The different coloured strings show different interactions.

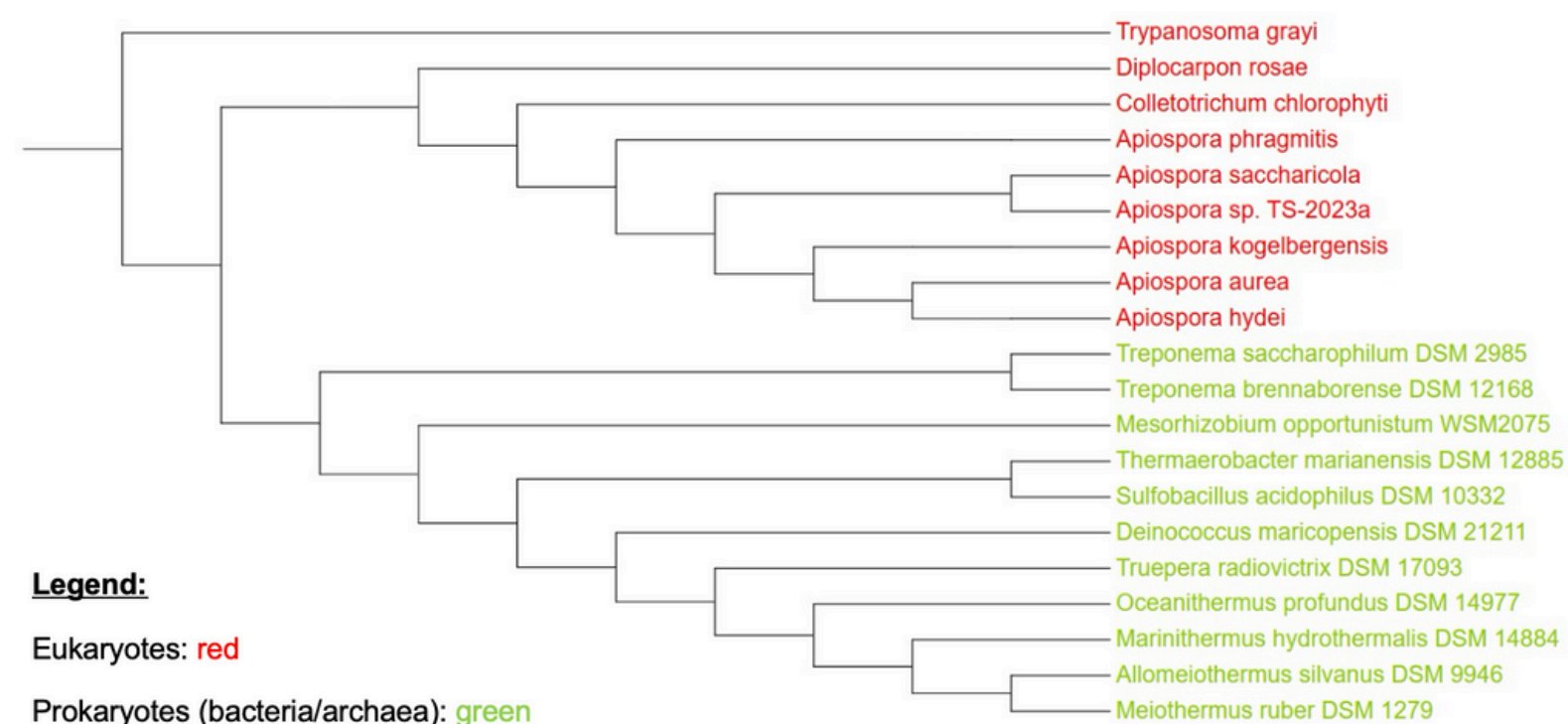


Figure 3: Phylogenetic tree results of the organisms who share the protein Arginase. Coloured by their taxonomy

From Interpro, the following information on the protein retrieved:

- Domains:**
 - Arginase/deacetylase domain (superfamily SSF52768)
 - Ureohydrolase domain
- Biological processes:**
 - arginine metabolic process → The chemical reactions and pathways involving arginine, 2-amino-5-(carbamimidamido)pentanoic acid.
 - L-arginine + h₂O > urea + L-ornithine
- Molecular function:**
 - arginase activity
 - metal ion binding → manganese ion binding
- Families:**
 - Gene locus:-Located at the rocF gene
 - rocF-arginase → ureohydrolase superfamily

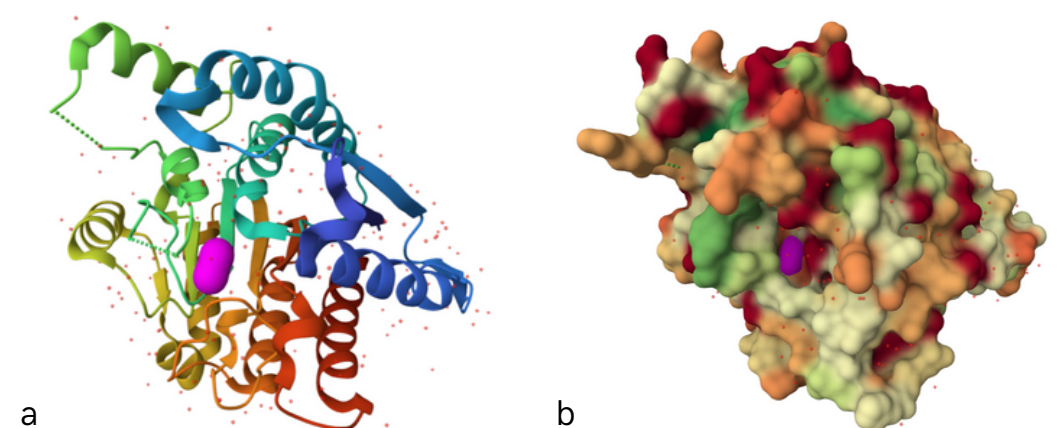


Figure 4: 3D structure (a) and molecular surface (b) representation generated with PDB of the protein Arginase.

REFERENCES

- Öztekin, M., Yılmaz, B., Ağagündüz, D., & Capasso, R. (2021). Overview of *Helicobacter pylori* Infection: Clinical Features, Treatment, and Nutritional Aspects. *Diseases*, 9(4), 66. <https://doi.org/10.3390/diseases9040066>
- H. Pylori infection*. (2025, 25 september). Cleveland Clinic. <https://my.clevelandclinic.org/health/diseases/21463-h-pylori-infection>
- Public Health Agency of Canada. (2024, 25 november). *Helicobacter pylori: Infectious substances pathogen safety data sheet*. Canada.ca. <https://www.canada.ca/en/public-health/services/laboratory-biosafety-biosecurity/pathogen-safety-data-sheets-risk-assessment/helicobacter-pylori.html>
- Zhang, J., Zhang, X., Wu, C., Lu, D., Guo, G., Mao, X., Zhang, Y., Wang, D., Li, D., & Zou, Q. (2011). Expression, Purification and Characterization of Arginase from *Helicobacter pylori* in Its Apo Form. *PLoS ONE*, 6(10), e26205. <https://doi.org/10.1371/journal.pone.0026205>
- Gobert, A. P., McGee, D. J., Akhtar, M., Mendz, G. L., Newton, J. C., Cheng, Y., Mobley, H. L. T., & Wilson, K. T. (2001). *Helicobacter pylori* arginase inhibits nitric oxide production by eukaryotic cells: A strategy for bacterial survival. *Proceedings Of The National Academy Of Sciences*, 98(24), 13844-13849. <https://doi.org/10.1073/pnas.241443798>
- Azizian, H., Bahrami, H., Pasalar, P., & Amanlou, M. (2010). Molecular modeling of *Helicobacter pylori* arginase and the inhibitor coordination interactions. *Journal Of Molecular Graphics And Modelling*, 28(7), 626-635. <https://www.sciencedirect.com/science/article/abs/pii/S1093326309001855>

CONCLUSIONS

The analysis shows that arginase is a **highly conserved** protein found in both prokaryotes and eukaryotes, indicating its essential functional role (figure 3). In *Helicobacter pylori*, arginase belongs to the **rocF family** and is part of the **ureohydrolase superfamily**. It converts L-arginine into urea and L-ornithine, reducing host nitric oxide production and **aiding immune evasion**. The enzyme is manganese-dependent and has a well-defined catalytic pocket formed by alternating α -helices and β -sheets. Encoded by the RocF gene at the **HP_1399 locus**, it is active in the cytoplasm. STRING analysis further suggests that arginase is part of a broader metabolic network crucial for *H. pylori* survival (figure 2).